PB86-216744

Forecasting Onsite Soil Absorption System Failure Rates

Urban Systems Research and Engineering, Inc.
Cambridge, MA

Prepared for

Environmental Protection Agency, Cincinnati, OH

Jul 86

U.S. Department of Commerce
National Technical Information Service
NTIS

PB86-216744

EPA/600/2-86/060
July 1986

FORECASTING ONSITE SOIL ABSORPTION SYSTEM FAILURE RATES

by

James Hudson
Urban Systems Research & Engineering, Inc.
Cambridge, Massachusetts 02138

Contract No. 68-03-3057

Project Officer

James F. Kreissl
Wastewater Research Division
Water Engineering Research Laboratory
Cincinnati, Ohio 45268

WATER ENGINEERING RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
CINCINNATI, OHIO 45268

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA/600/2-86/060	2.
3. RECIPIENT'S ACCESSION NO. PB86-216744/AS	
4. TITLE AND SUBTITLE Forecasting Onsite Soil Absorption System Failure Rates	5. REPORT DATE July 1986
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) James Hudson	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Urban Systems Research & Engineering, Inc. 1218 Massachusetts Avenue Cambridge, MA 02138	10. PROGRAM ELEMENT NO.
	11. CONTRACT/GRANT NO. 68-03-3057
12. SPONSORING AGENCY NAME AND ADDRESS Water Engineering Research Laboratory--Cinti, OH Office of Research and Development U.S. Environmental Protection Agency Cincinnati, OH 45268	13. TYPE OF REPORT AND PERIOD COVERED Complete
	14. SPONSORING AGENCY CODE EPA/600/14

15. SUPPLEMENTARY NOTES

James F. Kreissl Telephone: (513) 569-7611 FTS: 684-7611

16. ABSTRACT

A review was made of four previous statistical studies of onsite soil absorption system failure rates in four different areas of the United States. From this data base a series of forecasting techniques were developed with varying levels of accuracy for use by facility planners in evaluating comparative costs of alternative plans for pollution abatement in small communities. The required input data and expected level of accuracy to be expected for each forecasting method is also provided along with the assumptions made in development and testing of the models against available data.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT Release to Public	19. SECURITY CLASS *(This Report)* Unclassified	21. NO. OF PAGES 64
	20. SECURITY CLASS *(This page)* Unclassified	22. PRICE $11.95

EPA Form 2220-1 (9-73)

DISCLAIMER

The information in this document has been funded wholly by the United States Environmental Protection Agency under EPA Contract No. 68-03-3057 with Urban Systems Research and Engineering, Inc. It has been subjected to the Agency's peer and administrative review and it has been approved for publication as an EPA document.

FOREWORD

The U.S. Environmental Protection Agency is charged by Congress with protecting the Nation's land, air, and water systems. Under a mandate of national environmental laws, the agency strives to formulate and implement actions leading to a compatible balance between human activities and the ability of natural systems to support and nurture life. The Clean Water Act, the Safe Drinking Water Act, and the Toxics Substances Control Act are three of the major congressional laws that provide the framework for restoring and maintaining the integrity of our Nation's water, for preserving and enhancing the water we drink, and for protecting the environment from toxic substances. These laws direct the EPA to perform research to define our environmental problems, measure the impacts, and search for solutions.

The Water Engineering Research Laboratory is that component of EPA's Research and Development program concerned with preventing, treating, and managing municipal and industrial wastewater discharges; establishing practices to control and remove contaminants from drinking water and to prevent its deterioration during storage and distribution; and assessing the nature and controllability of releases of toxic substances to the air, water, and land from manufacturing processes and subsequent product uses. This publication is one of the products of that research and provides a vital communication link between the researcher and the user community.

This report describes a simple statistical forecasting tool which permits facility planners to realistically evaluate the onsite wastewater treatment alternative in small rural communities against centralized collection and treatment alternatives. The approach described requires only that there be documented failure rates for a sufficient period prior to the planning effort. Heretofore, only gross assumptions of failure rates during the planning period were used to compute present worth and annual costs of the onsite alternative.

Francis T. Mayo
Water Engineering Research Laboratory

ABSTRACT

Facility plans for small communities usually involve the continued use of many existing onsite systems. Almost all systems left in continued use will, eventually, fail -- from a variety of causes. Repair or replacement of these failing systems will be expensive. Such plans may also propose new onsite systems, which must be evaluated.

Local planners need to be able to forecast future failure rates for two reasons:

- To compare the costs and requirements for alternative methods of wastewater disposal, and
- To evaluate the effects of changes in onsite system codes, site evaluation tests, designs, installation practices, and management.

This paper provides a simple statistical forecasting tool useful at the local level. It builds on the past research in this subject area, and was tested using data from four areas of the country: Connecticut, Virginia, Massachusetts, and California.

Given an estimate of failure rates, it is easy to implement a forecasting system for future failures, and for the costs of those failures. Essentially, a certain fraction of all systems may fail in each year. A simple computer program, easily adapted for hand calculator, was developed to prepare forecasts of future failures. Given an estimated replacement cost, present values can also be calculated.

Data required include past failure experience, and information needed to divide the households in the community into cohorts (years of major code or practice revisions, counts of new systems installed each year, etc.)

Given those data, and the techniques presented in this paper, any moderate size community will be able to develop estimates of future failure rates and costs, and -- given moderate assumptions about effectiveness -- analyze the costs and benefits of changes in codes and practices. In addition, the model provides important inputs into the evaluation of changes in system configuration, such as a shift from onsite to centralized collection and treatment.

This report was submitted in partial fulfillment of EPA Contract No. 68-03-3057 by Urban Systems Research and Engineering, Inc. under the sponsorship of the U.S. Environmental Protection Agency. This report covers the period from 1981 through 1983, and the work was completed as of September, 1983.

CONTENTS

Foreword iii
Abstract iv
Tables vi
Abbreviations and Symbols vii
Acknowledgements viii

1. Introduction 1
2. Conclusions and Recommendations 4
3. Theoretical Background and Past Research 5
4. Summary of Existing Data 14
5. Analysis Results 18
6. Using the Technique 28

References 44
Appendices

A. Sample analysis 46
B. Student's t - Distribution. 54

TABLES

Number		Page
1	Basic regression results	19
2	Raw failure rates	21
3	Analysis of individual years	25
4	Temporal shifts in low percent failure rates	26
5	Comparison of predicted and actual failures	27
6	Table of failures	35
7	Table of survivors	36
8	Example summary failure rate presentation	37
9	Example problem tabulation	42
10	Overall results of example problem	43

ABBREVIATIONS AND SYMBOLS

$A_{i,t}$ = number of systems abandoned from cohort i in year t

b_i = coefficients estimated from data

B_0 = estimated parameter

B_1 = estimated parameter

$CP(t)$ = cumulative probability of failure through year t

f_i = failure rate for cohort i

F_t = number of failures

F_{rt} = number of reported failures which started in year t

F_{ut} = number of unreported failures which started in year t

$F_{i,j,t}$ = number of failures in cohort j for year t which are rebuilt as cohort i

N_t = number of systems still in service after t years

N_0 = number of systems constructed in year 0

$N_{i,t}$ = number of systems constructed in year t in cohort i

P = the overall failure rate, probability that any particular system will fail in any particular year

P_1 = failure rate for initial period

P_2 = failure rate for later period

R = fraction of unreported failures

$S_{i,t}$ = stock of systems in cohort i at the end of year t

T = age

X' = estimated parameter, correction factor for the effect of early failures

Y_1 = total years, initial period

Y_2 = total years, later period

ACKNOWLEDGEMENTS

The key acknowledgement for this study is to those who came before: the researchers who collected and analyzed data on failure rates of onsite systems. In particular, four groups receive our heartfelt thanks: Hill and Frink, for their Connecticut work; Clayton and Bowman for their extensive data collection in Fairfax County; Brian Scanlon for his work in Los Angeles; and Saxton & Zeneski for their Acton work. Without their data collection efforts, and their help in obtaining the raw data and reports, the analysis would have been impossible.

In addition, Steve Calichman helped us obtain the Acton data, and Gary Saxton's assistance was invaluable in designing the methods and performing the analysis. At USR&E, Patricia Deese and Cynthia Ernst provided many useful comments and research support. Helpful comments and encouragement were received from Bob Bowker, Jim Cox, and Jim Kreissl at EPA, and Mrs. Kathryn Wiker at EPA provided the final clerical effort.

SECTION 1

INTRODUCTION

Facility plans in small communities generally involve the continued use of many existing onsite wastewater treatment and disposal systems.

In evaluating the costs of various alternatives, it is important to forecast how many onsite system failures will occur each year, and how much effort will be involved for their replacement or rehabilitation. This report describes a simple forecasting tool for estimating the number of onsite system failures, given data on past failures, building activity, and the number of systems.

Developement of the forecasting tool was the major focus of this task, though not the only one. Differences in soil conditions, sanitary codes, construction practices, and maintenance will affect the useful life of onsite systems. The analysis of four communities presented here provides some guidance on analysis of useful lives. The model described in this report may provide a useful approach for field studies.

TYPES OF FORECASTING TOOLS

One of the most difficult problems in a method-developement study such as this is to find the right level for the analysis. The technique must use data which can be collected economically and must be within the abilities of the decision-making staff.

Numerous factors can affect whether a particular system will fail on a particular site. Many of these factors can be abstracted to the community as a whole and measured on an average, individual site, or aggregate basis. If a factor can be shown to be measurable, and the measured values have an effect on failure rates, then the factor should be included in a forecasting tool.

The basic approach taken here was to analyze existing data to determine which of the measured factors could be shown to have a significant effect on the failure rates, and then to develop an approach which used the key factors. Fortunately, the number of factors was found to be small, and the techniques which use them, straightforward. The data required are also limited, though these data may not be available in many communities.

There is a spectrum of techniques which could be used for forecasting failure rates. The method recommended here, and described in some detail in Section 6, is mid-range in complexity and data requirements. The various

techniques are:

1. Old is unacceptable: If a system is already in operation (or, perhaps, older than T years), it must be replaced. This is simple, but it may throw away effective systems which are currently in use.

2. Fixed Planning Horizon: Assume that every system lasts a fixed amount of time. Compute the number installed each year in the past, and assume that number will be replaced in each year when the time runs out. For example, assume a wastewater project is implemented in 1985, and that a 20-year life is accepted for past systems. Then all systems from 1965 or before will be replaced in 1985, all 1966 systems in 1986, all 1967 systems in 1987, and so on. This is simple to implement, and the counts can be found from building permits, but it will be independent of the existing failures or the failure rates. It will also ignore improvements made by better codes.

3. Raw failure rates: If the current data show, for example, 30 failures per 1000 households per year, then you could assume that 3% will fail in each future year. This will be more accurate than the first two methods, but does not consider the effects of aging systems or of changes in the codes, among other factors.

4. Cohort raw failure rates (the recommended approach): Divide the systems into groups by what set of codes were in effect when they were built and, if possible, other factors such as soils. Compute the raw failure rates for each cohort. Estimate future survivors in each cohort each year, and apply those failure rates to get annual failures. This requires more data, and is only as good as the cohorts you define. It ignores some timing issues, but it is the simplest method which allows consideration of changes in codes or operating practices, and differences in soils or other conditions.

5. Time-dependent cohort failure rates: Similar to method #4, but used different failure rates for early failures (first year or perhaps more), mature systems, and perhaps old systems. In most cases, the gains in accuracy are small and the extra work in developing the estimates is considerable.

6. Effects of construction year: Systems built in some years (particularly dry or wet, for example) may have different failure rates than systems built in other years. This can be measured using statistical techniques, and could be added to the analysis. Again, this is not worth the effort except in special cases. It can require considerably more data.

7. Full statistical analysis: There are numerous types of models which could be applied to this problem, given exceptional data. This would require the participation of statisticians or other

experts, but might result in somewhat improved forecasts of future failure rates.

The recommended method avoids the clear problems with the first two approaches, while providing most of the benefit of the later methods -- at least in the test examples shown in the reaminder of the report.

USE OF THIS REPORT

Users who are only interested in how to apply the method should skip ahead to section 6. Sections 3 - 5 provide the motivation for the approach, and tests against existing data.

SECTION 2

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

1. The best compromise between accuracy and usefulness in forecasting onsite soil system failure rates is the use of the cohort raw failure rate approach described and demonstrated herein because it allows the forecaster to evaluate code changes, soil differences and management method shifts.

2. The next best approach is the use of raw failure rates as an annual average during the planning period, but this method fails to account for major categories of changes or differences which can significantly influence future failure rates.

3. More complex methods may be feasible if sufficient data are available, but the cost of gathering these data and the complexity of the analyses are not generally justified for rural communities.

RECOMMENDATIONS

1. The use of rational failure forecasting methods should be applied to all small community facility planning alternatives which involve either continued use of existing or installation of new onsite wastewater disposal systems.

2. Application of the recommended failure forecasting method or more complex methods are recommended as a prerequisite to any community contemplating a major change in their onsite system codes to provide a basis both for that decision and for review of the effects of that revision at a later date.

SECTION 3

THEORETICAL BACKGROUND AND PAST RESEARCH

INTRODUCTION

Onsite wastewater disposal systems fail for many reasons. Failures may be caused by human errors such as poor construction practices, undersizing, overload by additional contributions or by the connection of extra appliances, etc. In attempting to develop a statistical model of the rate of failure for a community, these detailed causes for failure are not used. At best, the model may suggest that the system manager should expect "8 of every 100 systems to fail next year," based on past performance and the characteristics of those systems. In many situations, though, that is enough -- particularly if it can be inferred that a certain set of code changes could reduce the yearly failure rate.

This section presents a set of hypotheses about how systems fail, and how failures are reported. Given the types of data likely to be available in a community, it is possible to analyze how the failure rate should be affected by several critical variables. A discussion of theoretical model forms is also presented along with a review of past work on septic tank failure rates from the standpoint of the theoretical model and the goals of this project. Literature on this topic is limited, but the few key papers provide a great deal of useful information.

THEORY OF ONSITE SYSTEM FAILURE

Onsite systems can basically fail in three ways:

- they can fail to accept all of the wastewater received,
- they can fail to transfer all the wastewater to the soil, or
- they can fail to adequately treat the wastewater.

The first type of failure is identified by a backup of sewage into the house. The second causes ponding in the yard or short-circuiting to surface water; a treatment failure results in contamination of ground water.

In this report, we are concerned with the second type of failure, since treatment failures are often difficult to identify and trace.

Household backups (first type) generally are related to drainage piping or pretreatment system blockages, which bear little or no relationship to the soil absorption system performance. However, because such failures are often reported along with the second type of failure,

disaggregation can sometimes be difficult. This problem could be overcome with improved reporting procedures which differentiate the types of failure. Such improvements are to be encouraged.

Direct Causes for Failure

Surface failures can be caused by several problems within or outside the onsite disposal system:

1. Hydraulic overload through high household water use;
2. Solids overflow into the disposal field caused by lack of septic tank maintenance and pumping;
3. Scum overflow into the disposal field caused by a broken septic tank baffle;
4. Buildup of a clogging mat with insufficient capacity to accept the daily flow (caused by undersizing or overuse);
5. Reduced soil permeability caused by smearing in construction;
6. Poor distribution resulting in hydraulic overload of sections of the soil absorption system, leading eventually to insufficient capacity;
7. Low soil permeability.

Other causes are, of course, possible. Usually, an onsite inspection is necessary to determine the cause, or set of causes, of a particular failure.

Human Activity and Natural Conditions Related to Failures

The direct causes are not amenable to modeling. A more useful approach for this study is to analyze what human errors may have affected the life of the disposal site. The three basic types of errors are:

- Improper design for the conditions,
- Poor construction practices, and
- Insufficient maintenance.

Local and state codes attempt to control practices in each of these areas. The codes -- and their related enforcement mechanisms -- can be expected to affect the failure rate. Codes contain an assumed relationship among soil characteristics, size of dwelling, and size of the disposal system. Unfortunately, this set of assumptions is often inappropriate. For example, a code could adequately specify size for systems in sandy soils serving 3 bedroom houses, while inadequately sizing systems in tight soils serving 5 bedroom houses.

Given past experience and research, the following initial hypotheses should be tested:

1. Larger disposal fields will last longer in sandy soils, as long as the site characteristics (e.g. soils) and flow stay the same.

2. Tighter soils will typically lead to more failures, for a given size of system and load.

3. Greater water use will increase the chance of failure for any particular site and system size.

4. Adequate maintenance and pumping will extend the life of the system.

5. Greater control over construction (particularly increased field inspection) will decrease the chance of failure. This is especially true in fine-grained soils.

6. Use of alternating fields will decrease the rate of failure, for any given set of soil conditions, size, and loading rate.

7. Better distribution, particularly through pumping, will decrease the rate of failure.

8. There may be a high failure rate on some soils in the early stages of system operation.

9. With the exception of initial failures related to poor design or construction, systems are more likely to fail as they get older.

10. Failures may or may not be reported to the appropriate regulatory authority; some undoubtedly go unreported, no matter how strict the regulations.

A general model would consider all of these hypotheses, analyze whether they are supported by the data, and estimate their relative importance (based on the number of failures related to each). Any particular model must be evaluated based on its ability to analyze each of these hypotheses.

Model Structure and Functional Forms

The hypotheses listed above do not define one correct form for a model for predicting onsite failures. However, they are a starting point for consideration of functional forms, model structure, and availability of data. For example, consider the effects of system age. Consider a single onsite system which has been successfully operating for 10 years. What should affect the probability that the system will fail in its 11th year?

- Age might have an effect, with the probability of failure higher in the 16th year than in the 11th. This effect is expected to be small. In addition, <u>after the first few years</u>, the effect

of maintenance practices should average out, and the effect of construction practices should already be realized.

- o However, as the equipment in the systems ages, there should come a point where the failure rate rises sharply as components wear out. For conventional gravity-feed systems this may be well after construction, given the design life (e.g., 30-50 years) for pipes and tanks, but the point will come eventually. With more complex design, such as pressure distribution systems, this effect will be pronounced during earlier parts of the cycle.

- o Therefore, there may be a high rate of failure in the first few years (and often the first year) of operation, when the effects of poor construction are seen, and the system has not yet reached long-term equilibrium. This may also be the time of the first failures related to lack of periodic maintenance and pumping. There may also be a high failure rate at the end of the useful life of the systems. In between, the effect of aging is likely to be small.

If the effect of age is small, the probability of failure would not be correlated with system age, at least across a wide period (perhaps 5 to 30 years after construction). In that case, the population of systems in the middle age group will have a negative exponential form:

$$N_t = N_0(1-p)^t \qquad [1]$$

where

N_t is the number of systems still in service after t years

N_0 is the number of systems constructed in year 0

p is the failure rate, the probability that any particular system will fail in any particular year

This is a much simpler functional form than if the failure rate f depends on the age T, in which case the form is

$$N_t = N_0 \prod_{t=0}^{T} (1-p(t)) \qquad [2]$$

where $\prod_{t=0}^{T}$ is the product from year 0 to year T

Thus, one important test is to see whether the effects of system age are significant, independent of the effects of codes and other factors which change over time.

Other concerns can be handled similarly. For example, as discussed

above, it is reasonable to assume that the acute effects of construction practices (immediate failures caused by construction errors) will show up within the first five years of system operation - if not even earlier.[1] Also assume that reporting of failures will be relatively consistent over time, barring evidence to the contrary. That is, if 80% of all failures are reported in 1970, we will assume that 80% are reported in 1980.

# of reported failures = (1- R) (# of actual failures) [3]

where R is the fraction not reported, which is assumed to depend on location (or area), not time

On the other hand, there are certain processes which do not lead to simple functional forms. For example, the stress on a system is likely to depend on the flow, the system size, and the soil type, even given acceptable maintenance and equal construction practices. The best option is to consider all systems developed under equivalent codes and conditions to be a cohort. The model can then compare the performance of one cohort (one set of codes, soils, etc.) to another. It may be possible to say that one cohort has clearly better performance than another in a community, but might not be possible to identify exactly what aspects of that cohort cause the better performance.

Given the above discussion, the first step is to identify whether age has a major impact on the failure rate within a cohort. Then, it is possible to compare cohorts within a community. For example, the analysis can compare sets of systems built under different codes or on different soils.

Finally, it may be possible to compare cohorts from different communities, though the number of other factors potentially affecting survival is much greater. For example, one community may have higher water usage per household than another, so that houses with the same number of bedrooms place a greater stress on their systems. Weather and climate may be important, as may building materials or practices, appliance saturation, and numerous other factors. Comparisons across communities are likely to be difficult and unconvincing, given the currently available data.

PAST STUDIES OF FAILURE RATES

This section describes four major studies of failure rates, which are also the sources for the data used here. Survivorship curves and other measures of long-term failures have been used since at least the early 1960's. For example, the PHS studies which led to the Manual of Septic Tank Practice considered long-term survival of systems under various

---

[1]There may also be chronic effects of poor construction on system efficiency. Consider a system with poor distribution because of elevation errors. It may operate, but could be effectively equivalent to a system half its size with good distribution.

design parameters and soil conditions, and collected failure data (see Appendix C). Winneberger and DeWalle have also suggested ways to develop survivorship curves. However, only in recent years have large data sets covering hundreds of systems become available.

Such data are necessary when the analysis must deal with failure rates on the order of 1%, if any confidence is to be placed in the results.[1] Data sets with a few thousand systems are desirable, particularly if the results are to be differentiated by year of installation, soil type, or other factors. Much of the early literature is interesting, but provides little help in estimating useful models for forecasting failures.

Therefore, this section focuses on four studies which have reasonably large data sets. Three of these are published and widely available. They are:

1. Hill, D.E. and C.R. Frink, "Septic System Longevity Increased by Improved Design." Journal WPCF 52(8): 2199-2203, August 1980.

   Hill, D.E. and C.R. Frink, "Longevity of Septic Systems in Connecticut Soils," Bulletin 74, Connecticut Agricultural Experiment Station, New Haven, June 1974.

2. Clayton, J.W. "An Analysis of Septic Tank Survival Data 1952-1972 in Fairfax County, Virginia" Journal of Environmental Health, 36:562-567, 1974.

   Bowman, J.O. "A Study of Septic Tank Soil Absorption Systems" in N.I. McClelland and J.L. Evans (ed.) Individual Onsite Wastewater Systems: Proceedings of the Seventh National Conference, 1980. Ann Arbor: National Sanitation Foundation, 1981. pp. 109-114.

   Bowman, J.O., "Reliability of Onsite Sewage Disposal Systems in Fairfax County, Virginia", Journal of Environmental Health, 44(5):249-252, March-April 1982.

3. Saxton, G.B. and J.M. Zeneski, "Prediction of Septic-System Failures," Journal of the Environmental Engineering Division, Proceedings ASCE 105(EE3):503-509, June 1979. The data presented in the paper were updated to 1981 by field research for this project.

---

[1] The standard deviation of a proportion varies as the square root of p(1-p)/N, where p is the probability of failure and N is the number of systems. If p is small (e.g. under 5%) and the analysis is concerned with small variations in p (±1%), then many cases are needed to estimate p with any level of confidence.

The fourth is an outgrowth of the 208 Project performed by the Southern California Association of Governments:

4. Department of County Engineer-Facilities, Los Angeles California. Waste Treatment Management for the Malibu/Topanga Area, Task 2-1D, Phase II: SCAG 208 Areawide Waste Treatment Management Planning, June, 1978. (Additional material as a supplement to 208 Final Report, March 1980).

All four studies involved data collection on system installation, characteristics, and failures in specific areas:

- o Glastonbury, Connecticut (Hill & Frink)
- o Fairfax County, Virginia (Clayton, Bowman)
- o Acton, Massachusetts (Saxton & Zeneski)
- o Malibu-Topanga area, Los Angeles, CA (Los Angeles County)

The analysis methods used in the four studies are discussed below; the available data are presented in Chapter 3.

Hill & Frink

The WPCF paper by Hill and Frink is an update and improvement on the earlier work, and uses some advanced statistical methods for the analysis. The earlier work demonstrated a normal distribution of failures around a mean, except for a set of early failures. This is a natural tendency for a large sample. The first paper estimated the cumulative percentage failure by linear regression on probits. The more recent work used a log-normal transformation with a special weighting to take account of the increased failure rate in the first few years. Analysis was made individually by soil type and age, but not by both at once. That is, they analyzed systems built since 1972 versus earlier systems, and systems built on compact glacial till versus stratified sand and gravel, but did not report analysis of pre- and post-1972 systems specifically on compact till.

The estimation procedure is reasonable, but appears overly complex. It involves estimating the cumulative probabilities of failure, rather than the probability density function. Typically, this will increase the apparent "fit" of results without adding any real explaining power. The assumed form is that the cumulative probability is:

$$CP(t) = B_0 + B_1 * \log(t+x') \qquad [4]$$

where CP(t) is the cumulative probability of failure through year t

B_0, B_1, and x'are estimated parameters (with x' specifically being a correction factor for the effect of early failures.)

The work by Hill & Frink is pioneering in both its data collection and methods, but simpler and clearer models can fit the data well, as presented in Section 5.

Bowman, Clayton

The three papers from Fairfax County report two analyses: one covering 1952-1972, and an update to 1978 with a somewhat different sample. They are primarily data collection studies, identifying the number of failed systems and the percentage failure over long periods.

Saxton & Zeneski

This study used data from 1968-1977 for Acton, Massachusetts, and divided systems into age cohorts. Soils data were not available. Cumulative failure rate curves were developed following the first Hill & Frink paper, and half-lives estimated for each age cohort. The authors then took a further step. First, they estimated the number of systems constructed each year from 1927 to the end of the planning period. They could then use the probability of failure to compute the expected number of failures in any future year. For example, if there were 100 systems from 1960 operating in 1980, and a 2% failure rate for 20-year old systems, then 2 failures of 1960 systems would be expected in 1980. By summing across all system ages, they could develop total expected failures per year; for example, there might be 83 expected failures in 1980, 72 in 1985, and 134 in 1990.

This is a simple simulation approach, but an important one: it allows the estimated failure rates to be converted into estimated numbers of future failures, which can then be converted to future costs for repair, rehabilitation, or replacement.

Los Angeles

The analysis in Los Angeles covered 8 areas with known problems, but only approximately 30 systems per area. The analysis included developing the distribution of system lives for both systems which had failed and all systems, and some regression analysis on the cumulative curve. There is little new in this analysis, but the important factors were considered. Unpublished work[1] analyzed the effects of soils types at an aggregate level.

Summary Evaluation

The past efforts have concentrated on statistical analyses of the cumulative failure curve, and have subdivided the data into either age or soils type cohorts where the data were available. An analysis of the density function is usually better for differentiating between models, and closer to the main goal of forecasting how many failures are likely to occur in year t, not how many have occurred through year t. In addition, all functional forms will tend towards a normal distribution around a

---

[1]Personal Communication with B. Scanlon, Department of the County Engineer, Los Angeles, September, 1981.

half-life (given a single population), and the specific functional form of the density function may be important. Therefore, the approach presented in Section 5 differs somewhat from these past efforts, though leading to similar results in most cases.

SECTION 4

SUMMARY OF EXISTING DATA

This chapter describes the available data from the four past studies, as updated under this project. It describes the data on systems and failures, presents data on relevant codes, and presents data on natural conditions. Note that section 6, describes the minimum and ideal data for facility planning applications of the models presented in this report.

SYSTEMS AND FAILURES

Glastonbury

The data in the updated Hill & Frink study covered 3,156 septic systems in the community. Failures were defined and dated through repair permits, and did not include mechanical failures. Certificates of Occupancy or Building Codes were used to determine the date of first occupancy and initial system use. Systems ranged from 1 to 31 years of age.

Fairfax

The 1978 Fairfax study reviewed records on 5,061 systems, and included both mechanical failures and malfunctions of the disposal systems. Like Glastonbury, reported failures were used. Systems ranged up to 26 years old, with age defined by installation date rather than occupancy.

Acton

The Acton study reviewed repair permits, but only those issued from 1968-1977. Both recorded data and the Health Director aided in determining which repairs were related to failures. Since all systems in the town were onsite, population growth was used as a surrogate for the number of systems constructed. Numerous pre-1943 systems were included. The repair permits and number of new systems were updated to mid-1981. The total number of failures was about 400.

Los Angeles

The Los Angeles study used random sampling of data on addresses in the eight study areas. Fifth addresses were initially used in each area, with data collection completed on approximately 30 structures in each. The data search started with use of building permits for age of structure and construction of replacement systems. Health Department files were also inspected for additional data, and for noted problems. Sewer permits

were assumed to be evidence of failure unless there was simultaneous addition or other building.

However, the various records were incomplete, and it was assumed that some failures might not be reported. Therefore, the records of pumping companies were checked to identify sites of frequent pumping, which were numerous. Normal maintenance pumping was rated by the pumpers as "practically nonexistent" in the area, though the study assumed that systems which were pumped only once were not failures. This study probably has the best data of the four on the number of actual (as opposed to reported) failures.

The study included a few apartments and commercial structures. Some of the dates of first construction or occupancy were unavailable, so that some ages are only known in the form "greater than x years." The analysis performed by the 208 project expanded the results to the full number of structures in each area.

CODES AND ENFORCEMENT

Glastonbury

The study reports two major changes in the codes affecting systems in Glastonbury.

> "A major revision in the health codes in 1961 ... virtually tripled the size of leaching fields in most soils" and
>
> "The year 1972 is significant because in that year the Glastonbury Health Department initiated the requirement of spring percolation testing and deep pit observation for all housing developments. This action was designed to evaluate more accurately the results of percolation tests taken at a time when percolation rates are normally slowest."

Therefore, the critical dates which might affect system performance are 1961 and 1972, and three age cohorts are used here: up to 1961, 1962-1971, and 1972 until now.

Fairfax

The critical dates in Fairfax County are less clear than in Glastonbury. The first major change in ordinance came in 1956, after high early failure rates and a period of fast growth. It was based on a soil survey, funded in 1953, and revised soil evaluation procedures; sizing stayed relatively unchanged. However, it has been reported[1] that installation practices and system sizing may not have been greatly affected until the early 1960's.

---

[1]Personal Communication, Fairfax County Health Department, September 1981.

A further set of revisions was passed in 1973, based in part on the results of the 1972 longevity study. The current code provides for alternating fields with diversion valves, and also requires a reserve area on each lot. Trench spacing was modified, but other design criteria were unchanged.

Acton

The only major change of significance in Acton which can be included in the analysis is a shift in practices in late 1971. At that point, both design and installation requirements were made more stringent. There were also some changes before the major growth occurred (c. 1950), but exact ages were unavailable for the individual structures. The best alternative was to use a 1943 USGS map to identify pre-existing structures at that date, and consider them to be a cohort. Therefore, the Acton results use three cohorts: pre-1943, 1943-1971, and 1972 until now.

Los Angeles

> "Detailed disposal field and seepage pit area requirements based on soil classifications and the number of bedrooms were first adopted in the 1948 ordinance.... In 1958, the disposal area requirements were revised. The disposal area requirement for drainfields (most of our systems) increased only slightly. However, the area requirements for seepage pits have increased by 120% to 67% depending on the soil type. These requirements have remained unchanged through the present..."[1]

According to the study director,[2] there have been no consistent changes in installation inspections, construction mehtods, record keeping, or the like during the period of concern. All structures are from 1948 or later. Therefore, there may be a difference between pre-1958 and post-1958 systems, but that is the only age cohort of interest.

PHYSICAL PARAMETERS

The major physical parameter of interest available for analyzing the performance of onsite systems is the soils type. No percolation rate test data or other direct measurements of supposed suitability were available, and only limited soils type information.

Glastonbury

The data from Glastonbury has the most soils information of the four locations. Soil types were determined from Soil Conservation Service survey sheets for each lot. Five types were used:

---

[1] Malibu/Topanga report, p. 12. (see references)

[2] Personal Communication, Brian Scanlon, September 1981.

- o Stratified sand and gravel,
- o Loose glacial till,
- o Compact glacial till,
- o Shallow to bedrock, and
- o Poorly drained.[1]

These define five cohorts for the analysis.

Fairfax

No soils data were available for the systems.

Acton

No soils data were available for the systems.

Los Angeles

Two of the eight study areas had consistent soils. The Beachfront areas are on sand, and the Point Dunne area is generally expansive soils with large amounts of clay.[2] The other areas have variable soils, and no identification was made of soil type for an individual lot. Several of the areas have steep slopes (e.g. Zuma Canyon, Coastal Hillsides). Maps are available, but there are no data on the slopes for any particular study lot.

---

[1]Hill & Frink, WPCF, op,cit., p. 2199.
[2]Personal Communication, Brian Scanlon, September 1981

SECTION 5

ANALYSIS RESULTS

This section presents the models developed in this analysis for each of the four study areas, shows the results of analyses of individual areas, and shows joint results.

RESULTS FOR INDIVIDUAL AREAS

The first analysis was performed using linear regression on 14 individual cohorts. The same model was used for each regression. If the dependent variable was the number of failures F_t over the number of systems N_t, a logistic transform to $\ln (F_t/(N_t - F_t)$ might yield an improved model, but the extra accuracy may not be worth the additional effort. The linear regression is:

$$\text{Failure Rate} = b_1 + b_2(\text{Age}) + b_3(\text{Early}) \qquad [5]$$

where

Age = age of system in years
Early = 1 for ages 1-5[1]
0 otherwise
b_n = coefficients estimated from the data.

The areas, numbers of years of recorded data, peak number of systems, coefficients, standard errors, and regression statistics are shown in Table 1.

These results show, at most, limited age effects. Only one of the age variables approaches statistical significance (Fairfax, 1963-1968), and it is of the wrong sign: it shows systems failing less often as they get older. There appears to be an early years effect in Glastonbury for stratified sand and gravel: these systems survive well in the early years, and move more often later.

However, the data demonstrate few age or soils effects besides those used to define cohorts. Therefore, Table 2 shows the raw failure rates for each cohort: the number of failures per year of system operation (e.g., 1% = failure per hundred system-years).

---

[1]Shorter "early" periods were also tested, with similar results.

TABLE 1. BASIC REGRESSION RESULTS

	# years	b_1	AGE b_2	EARLY b_3	R^{2}*
Fairfax County					
to 1962	26	.22	.01 (.008)	-.18 (.153)	.20
1963-1968	17	.42	-.024 (.012)	-.24 (.13)	.23
Glastonbury					
Hardpan	20	1.36	-.03 (.07)	+.69 (.90)	.13
Stratified	30	1.67	-.002 (.02)	-1.24 (.41)	.36
Till	22	2.07	-.02 (.03)	-.63 (.49)	.08
Bedrock	25	1.28	.009 (.03)	-.47 (.53)	.10
Wet	15	3.57	-.21 (.13)	-1.59 (1.16)	.19
Los Angeles					
Coastal Hillsides	15	5.5	-.14 (.34)	-3.3 (3.2)	.12
Zuma Canyon	15	-1.29	.44 (.43)	1.05 (4.06)	.15
Fernwood	15	1.45	.06 (.26)	-.93 (2.37)	.08
Civic Center	15	-1.46	.46 (.35)	1.98 (3.20)	.17
Beach Front	19	3.98	-.06 (.31)	-1.43 (3.89)	.01

TABLE 1 (continued)

	# years	b_1	AGE b_2	EARLY b_3	R^{2}*
Topanza	15	2.22	.15 (.35)	-2.03 (3.23)	.20
Pt. Dunne	17	5.52	-.18 (.22)	-3.54 (2.37)	.16

*R^2 is a measure of the explanatory power of the estimate based on the model selected. Standard errors in () below coefficients.

TABLE 2. RAW FAILURE RATES (failures/year of operation)

	System Years	Total Failures	% Failure Rate
Los Angeles			
Total	3799	87	2.3
Pt. Dunne	628	17	2.7
Topanza	390	8	2.1
Beach Front	613	19	3.1
Monte Nido	427	2	0.5
Civic Center	473	13	2.7
Fernwood	431	6	1.4
Zuma Canyon	337	7	2.1
Coastal Hillside	500	14	1.8
Fairfax			
Total	60,355	110	0.18
to 1962	32,270	80	0.25
1963-1968	17,455	30	0.17
1969+	10,630	0	0
Glastonbury			
Total	35,320	501	1.42
to 1962	23,250	370	1.59
1963-1971	9,773	121	1.24
1972-1978	2,297	9	0.39
Poorly drained	1,562	22	1.41
Shallow to bedrock	4,788	62	1.29
Loose till	6,332	110	1.74
Stratified sand & gravel	19,069	265	1.39
Compact till	3,569	42	1.18
Poorly drained to 1962	723	8	1.11
1963-1971	739	13	1.76
1972-1978	100	1	1.0
Shallow Bedrock to 1962	3,783	53	1.40
1963-1971	884	9	1.02
1972-1978	0	0	0
Loose Till to 1962	3,798	73	1.92
1963-1971	1,837	31	1.69
1972-1978	697	6	.86

TABLE 2 (continued)

	System Years	Total Failures	% Failure Rate
Stratified sand & gravel			
to 1962	12,911	214	1.66
1963-1971	5,016	49	0.98
1972-1978	1,142	2	0.18
Compact Till to 1962	2,473	27	1.09
1963-1971	1,007	15	1.49
1972-1978	89	0	0
Acton			
1944-1971	NA	NA	2.14
1971+	NA	NA	0.69

In general, newer codes lead to lower failure rates. Since this is the basic goal of improved codes, it is not surprising. In Glastonbury, systems built before 1971 are missing early years of data, which could affect this analysis. However, Fairfax and Acton confirm these results. All three areas show a major drop in failure rate for newer systems, probably related to changes in codes and installation policies.

There are also differences related to subareas, shown in Los Angeles and Glastonbury. However, these are harder to explain. In Los Angeles, the eight areas are distinct, but there are no soils data and the sample sizes are small (20-30 households). There may be a real longevity difference between "Beach Front" and "Monte Nido", but this is arguable.

In Glastonbury, the highest failure rate occurs on loose glacial till, and the lowest on compact till, overall. This follows the results presented by Hill & Frink, except that they also found the poorly drained soils to have long lives. However, that was also the smallest sample by far, with the least reliable results.

Normally, the poorly drained, shallow to bedrock, and compact till soils would be considered marginal for conventional onsite systems. The loose till would seem best. However, longevity depends on the match between the codes and the systems and soil characteristics. Systems on difficult soils, but with careful design and installation, may be more successful than underdesigned systems on well-drained soils.

One further statistical test is useful: to determine whether individual years within a cohort show failure rates significantly greater or worse than average. This analysis was performed by the 1963-1971 stratified cohort in Glastonbury, as shown in Table 3. Even assuming perfect independence between years for single system,[1] the probability of failure in the worst year is not significantly greater than the mean over the whole population.[2] There may be good and bad years for construction, but they do not appear significantly within this cohort. The Glastonbury data showed significant differences between the age cohorts with different codes, but not between years within a cohort. There is some difficulty in defining the appropriate sample size, which affects what results are statistically significant. However, the cohorts differ significantly in failure rate.

COMBINED ANALYSIS

The previous section estimated and modeled failure rates for 28 cohorts from 4 areas. Differences were apparent between cohorts, and

---

[1] i.e., the probability of failure for system X in year t is not dependent of the probability in year T.

[2] The difference is less than $2\sqrt{P\ (1-P)/N}$ when p = .0166 and N = 785. If we assume perfect dependence, then N = 54 and the difference is much less significant.

these seem related to the definitions of the cohorts. The simple failure rates appear appropriate in most circumstances for forecasting.

Use of the single failure rates is necessary for the forecasting approach described in Section 6. However, the question arises as to whether more can be said about the differences between the failure rates for the various cohorts.

Table 4 presents failure rates of temporal cohorts. Fairfax has a more restrictive code than the other areas, plus the requirement for an alternating field. The code appears to have been effective. Note that approximately half of the total Fairfax failures were mechanical problems (e.g. broken plumbing) not reported in the data set, so the total failure rate is roughly double the amount shown in the table. Both the earlier and later Acton and Glastonbury codes fall between Los Angeles and Fairfax County in both absorption area and inspections, as well as in failure rates.

At this point, the data are too limited to determine why community failure rates differ. Soils, codes inspection, construction and maintentance practices, usage, and design all vary. Larger systems appear to fail less frequently, all else being equal. Stricter codes seem to work, in the areas studied. As Hill & Frink determined, Glastonbury should probably increase sizing and/or tighten construction inspection for loose glacial till relative to other soils. There is some evidence that size is a key factor, with larger systems almost universally leading to fewer failures. However, data for more areas would be needed to confirm this.

There are few surprises at the community scale, if only because data are available on too few communities. For individual cohorts, though, there appear to be fewer "early" failures than expected, and no direct effect of age over the first 30 years. The forecasting approach described in Section 6 has been developed on this basis.

VALIDATION

The only true validation would come from forecasting future rates of failure. However, a partial validation can be done using the available data. For the Glastonbury case, for all pre-1962 systems, the failure data to 1972 were used to estimate a raw failure rate of 2%. Then, this resulting estimate was used to generate "forecasts" for 1973-1977. These are compared with actual failures in Table 5.

The table shows some large differences, especially in 1973, but overall agreement. The model seriously underpredicts 1973 failures, and overpredicts in each other year. There is no obvious explanation.

TABLE 3. ANALYSIS OF INDIVIDUAL YEARS (Glastonbury/Stratified)

	Starting # Systems	A System Years	B Failures	C % Failures
1963	54	785	13	1.66
1964	54	753	8	1.06
1965	40	543	5	.92
1966	46	567	6	1.06
1967	52	596	4	.67
1968	71	754	6	.80
1969	44	431	2	.46
1970	36	314	2	.64
1971	36	273	3	1.10
Total	433	5,016	49	0.98

TABLE 4. TEMPORAL SHIFTS IN LOW PERCENT FAILURE RATES

System	Installation Dates	% Failures Per System Year of Operation
Fairfax	to 1962	0.25
	1963 - 1968	0.17
	1969 +	0
Acton	1944 - 1971	2.14
	1971 +	0.69
Glastonbury	to 1962	1.59
	1963 - 1971	1.24
	1972 - 1978	0.39
Los Angeles	All years	2.3

TABLE 5. COMPARISON OF PREDICTED AND ACTUAL FAILURES[1] (Glastonbury, pre-1963)

	Observed System Years	Predicted Failures	Actual Failures	Difference	Cumulative Difference
1973	984	19.7	32	-12.3	-12.3
1974	951	19.0	18	1.0	-11.3
1975	926	18.5	13	5.5	- 5.8
1976	912	18.2	15	3.2	- 2.6
1977	896	17.9	15	2.9	+ 0.3

[1]Failure rates of 2.0% calculated on data through 1972.

SECTION 6

USING THE TECHNIQUE

The models presented in the previous section were used to estimate failure rates for onsite systems, based on data about past performance. For facility planning, that is only part of the process. This section discusses the use of these techniques in facility planning, including suggestions on data sources, methods, and interpretation of the results.

This section describes the process, including data collection, estimation of failure rates, simulation of projected failures and costing of future repairs or replacements for failures. Examples are provided where appropriate.

THE GENERAL PROCESS

If one knew the number of failures of existing systems which would occur in each future year, and the cost of repair for each of those failures, then it would be easy to include those costs in facility planning. The following process attempts to compute those two items.

For facility planning, cost per failure is fairly easy to estimate. For example, a cost per replacement in constant dollars may be included in the plan already. If not, then current costs for an average replacement can be used, with the assumption that these will stay the same (in constant dollars) in the future.

It is more difficult to predict the number of failures for each year of the planning period. That requires knowledge of the number of systems currently operating (the "stock"), the number of new systems being constructed in each planning year and area, and the failure rates for each cohort.

Step 1 is to divide the current systems into reasonable cohorts. Each cohort is a group of systems with similar conditions and similar probabilities of failure. Cohorts may be defined by soils conditions, age, etc., but there should be less than 20 for almost any community. In the four areas studied, Glastonbury had the most: three age classes and five soils types, totaling 15 cohorts. Acton and Fairfax each had three groups; Los Angeles had 8. The number of cohorts and their definition depend on the available data, but cohorts should be based on well-defined dividing points such as differences in soil type or in the codes affecting construction.

Set up cohorts which might show different failure rates. However, they must use data which is possible to collect; every system in the planning area must be assigned to a cohort. Small cohorts -- under 500 system years -- are likely to give unreliable results.[1]

Step 2 occurs once the cohorts are defined. It is then necessary to determine how many systems are in each cohort, and how many newly constructed systems will be added to each. For example, if the plan is being prepared in 1980 and a new code was introduced in 1976, it would be reasonable to assume that all new construction will follow that code. Therefore, all new systems will be in the 1976 and later cohort. Sometimes, it may be necessary to divide the future growth into classes (e.g. by soils type or dwelling size, if those define cohorts) and to forecast growth in each class. For example, the projections might include 150 homes on sand and 57 on clay in 1985. Population projections must be converted to changes in the number of systems in each cohort in each planning year. Step 2 describes the cohorts and projection of growth. It includes a discussion of what to do about systems which are no longer used, even though they did not fail (e.g. systems replaced by sewers, systems for houses which were destroyed by fire, etc.).

Step 3 covers the estimation of failure rates for cohorts, using the methods described and tested in the previous section. Here, the major concern is available data on past failure rates. In addition, it is important to know what to do when data simply is not available -- how to use limited data to make the best possible guesses of failure rates, and how to vary the failure rates in later analysis to cover a reasonable range of possible values.

Finally, step 4 presents an approach for taking the data on failure rates and on the stock of systems in each cohort, and developing estimates of the failures from each cohort in each planning year. An example is included.

STEP 1: DEFINING COHORTS

The number of cohorts depends on the amount of data and the amount of effort available for analyzing the data. The two suggested ways to define cohorts are using codes and using soil types, but other factors could be significant in some cases.

Consider codes first. At some points in the past, the community probably made major changes in onsite system design and installation codes. Septic tank size, absorption area, allowable types of disposal systems (trenches, pits, etc.), and site classification may all have changed.

---

[1]For example, with 500 independent observations and a failure rate estimated at 2%, the standard deviation is 0.7%. Since the systems are not all independent, the true standard deviation would be higher.

In general, start with the assumption that any major change affecting the sizes, types, locations, or construction of onsite systems may affect the future failures of those systems. For example, if percolation tests were used up to 1970, and inspection pits ever since, systems built recently may survive longer than the older ones (or vice versa).

It should be fairly easy to determine the dates of major changes in codes. Then, define the cohorts as systems built between one such change and the next. For example, the Glastonbury codes were revised in 1961, and again in 1971. There are three cohorts:

1. Systems built up to 1961.

2. Systems built 1961-1971.

3. Systems built since 1972.

A typical community will probably have 2-4 critical dates within the last 30 years, and therefore 3-5 cohorts.

Now consider soils data. If there are adequate soils maps, or information on the site plans, it may be possible to classify which systems are on what types of soils. This was done in Glastonbury, using five types from Soil Conservation Service Maps:

Stratified sand and gravel,

Loose glacial till,

Compact glacial till,

Shallow to bedrock, and

Poorly drained.

There may be better data in some communities. Good soils data, or other data covering physical conditions (e.g. site test results), should be used to define cohorts. A community might, for example, have codes which require different system sizes for percolation test results as follows:

1. Less than 1 inch/hour

2. 1-10 inches/hour

3. etc.

If the waste water managers are interested in how well the test reflects reality, they could set up cohorts for each of these categories, and see if the failure rates vary from category to category.

Usually, there will be one cohort for each combination of age class and physical condition (for example, "systems built on sand 1957-1972").

However, that will not always be the case. For example, if soils information has only been recorded since 1973, but there was a code change in 1957, the following cohorts might be used:

- o Systems built through 1957
- o Systems built 1957-1972
- o Systems built on clay, 1973-present
- o Systems built on till, 1973-present
- o etc.

STEP 2: DATA DEVELOPMENT

First, gather data on the existing systems in the community. For each system, collect the following:

- o Date of installation
- o Date of relevant SAS failure (if any)
- o Any other information needed to classify this system into one of the cohorts (soils type, test results, number of bedrooms, etc.)

That is all the data needed for the current systems. However, the model also requires forecasts for the planning period. Those forecasts are of two types:

- o Number of new systems built for each cohort each year, and
- o Number of systems abandoned for each cohort each year.

The first part - estimating new construction - is relatively easy. The other parts of the planning process should provide the projected total of new constsruction each year. These need only to be allocated to the various cohorts (for example, estimate the number on sand, clay, etc.). If there will be a major change in codes, use it; otherwise, assume the new construction is built according to the current code.

There are two types of abandoned systems: systems which are replaced by sewering, and systems which serve houses which are destroyed or abandoned. Given any set of sewer plans, determine which systems will be replaced by sewer connections, and approximately when. If, for example, Forest Street will be connected to the sewers in 1987, and there are 20 houses on Forest Street, determine which houses are in what cohort, and list them. For example, sewering Forest Street might lead to the following changes:

- o Abandon 5 pre-1958 systems on sand

- o Abandon 4 1958 systems on sand/clay
- o Abandon 7 1958-1972 systems on sand
- o Abandon 2 1958-1972 systems on sand/clay
- o Abandon 2 1973 or later systems on sand

Those changes would affect the number of systems in each cohort in 1987.

Fires and other problems might also lead to systems being abandoned. Collect data on the abandonment rate from the local planning agency, or from insurance companies or builders.

For example, suppose that.08% of the houses in the community are destroyed or abandoned in an average year. Then, for each cohort, make the simplest assumption: assume that .08% of the houses in that cohort existing in the starting year are abandoned each future year. If there are under 1000 houses in the service area, ignore this problem -- the effect is very small.

This data development requires that each system be allocated to a cohort, using its age and other data. Failures must be identified from community records or, perhaps, pumper data. And, the planner must estimate future construction in each cohort, and replacement by sewers or abandonment.

There is one further data item, which is difficult to determine: the fraction of failures which go unreported. A community survey can estimate the fraction of all systems which are unreported failures at a particular moment. However, this is not enough -- the analysis requires the fraction of failures which are not reported. Professional judgement by code enforcement officials is one source for such estimates. Or, if data can be collected on the ages of unreported failures, then the factor can be estimated:

$$R = \sum_{t=1}^{I} \frac{F_{ut}}{F_{rt} + F_{ut}} \qquad [6]$$

where

R = fraction of unreported failures

F_{ut} = # of unreported failures which started in year t

F_{rt} = # of reported failures which started in year t

Note that the number of total failures is simply $F_{rt} + F_{ut}$, or $F_{rt}/(1-R)$. If 100 failures are reported, and we expect 1/3 of the failures to go unreported, then there should be 150 actural failures.

The model reported failures for future year. Given an estimate of R, the plan can forecast total failures. The reported failures are simply multiplied by (1/1-R).

STEP 3: ESTIMATING FAILURE RATES

This step only uses the data on the existing systems, not the forecasts. It requires organizing the data in one way to test for early failure rates, and in another way for estimating overall failure rates for the cohort. The analysis described below is repeated for each cohort.

Early (and Late) Failure Rates

The goal is to determine how many systems failed in their n^{th} year of life, and how many completed n years. This must be done for the full cohort. It is a three part process.

Take all the systems in the cohort. The number which successfully completed each year of operation depends on how old they are, and how many failed. For example, assume the systems were built as follows:

- 10 in 1975
- 8 in 1976
- 17 in 1977
- 22 in 1978
- 5 in 1979
- 12 in 1980
- 22 in 1981

Now state when the failures occured, for each year of installation. For example, take the systems built in 1978:

- 3 systems failed in their first year of operation
- 0 systems failed in their second year of operation
- 2 systems failed in their third year of operation
- all the others completed their third year.

Set up a table similar to Table 6, and use that table to create another similar to Table 7 by substracting down each column. For example, 22 systems were built in 1978. Three failed in their first year, so 19 were working after one year. None failed in their second year, so 19 were still working after two years. Then two failed, and 17 were working after three years. Table 6 shows the failures each year; Table 7 presents the

number still working.

Now compute failure rates for each year of life, as follows: take the total number of failures in a year, and divide by the number of systems surviving to the start of that year. For example, 4 systems failed in their third year out of the 46 systems still working at the end of the second year. That gives a failure rate of 8.7% per year, for systems 3 years old.

The result is a table with three columns, similar to Table 8: the number of systems surviving to year T, the number of failures in year T, and the failure rate for year T. The first two columns are the new totals of failures and survivors. The next question is whether there is a significant difference between early failure rates and later failure rates. That requires some calculation and a set of statistical tables. First, see if there is some age at which the failure rates appear to change dramatically. In the example, there seems to be a difference between the first three years of life, and the later years. Compute the average for each group, as follows:

Early group	Total years = 68+55+46 =	169
	Failures = 6+3+4 =	13
	Failure rate = 13/169 =	7.7%
Late group	Total years = 28+13+7 =	48
	Failures = 1+2+0 =	3
	Failure rate = 3/48	6.3%

TABLE 6. TABLE OF FAILURES

	Year Built						
	1975	1976	1977	1978	1979	1980	TOTAL
Failures in each year of operation							
year 1	1	0	0	3	0	2	6
year 2	1	0	2	0	0		3
year 3	0	1	1	2			4
year 4	0	0	1				1
year 5	1	1					2
year 6	0						0

TABLE 7. TABLE OF SURVIVORS

	Year Built						
	1975	1976	1977	1978	1979	1980	TOTAL
Survivors to start of each year of operation							
number built	10	8	17	22	5	12	74
year 1	9	8	17	19	5	10	68
year 2	8	8	15	19	5		55
year 3	8	7	14	17			46
year 4	8	7	13				28
year 5	7	6					13
year 6	7						7

TABLE 8. EXAMPLE SUMMARY FAILURE RATE PRESENTATION

Year	Systems at Start	Failures	% Failure Rate
1	68	6	8.8
2	55	3	5.5
3	46	4	8.7
4	28	1	3.6
5	13	2	15.4
6	7	0	0.8

The statistic is the test for a difference in proportions, and we will assume a single underlying population (the same households). The formula is:

$$z = [P_1 - P_2]/[(pq)^{1/2}x] \qquad [7]$$

where

P_1 = failure rate, initial period

P_2 = failure rate, later period

Y_1 = total years, initial period

Y_2 = total years, later period

$p = (Y_1P_1 + Y_2P_2)/(Y_1 + Y_2)$, the overall failure rate

$q = 1 - p$

$x = [(Y_1 + Y_2)/Y_1Y_2]^{1/2}$

In the example, N_1=169; N_2=48; p_1=0.077; p_2=0.063. Thus, p = 0.074, q = 0.926, x = 0.16 and the statistic value z is 0.33. There are N_1+N_2-2 degrees of freedom, (215) and the statistic fits the Student's t-distribution (see Appendix B). The tables show that a value of roughly 2.0 is necessary for 95% confidence that the failure rate in the first three years is different from the rate in the later years. Therefore, these data could easily have been caused by chance failures. The same analysis could be done for one or two years, or potentially by months.

The analyst needs to be at least 95% confident that the failure rates are different for it to be worth the extra effort of using two rates in forecasting. That means a value of the t-statistic greater than the value in the 0.05 column of the tables. Given the analysis in Section 5, this should be uncommon.

Computing Overall Failure Rates

If the data have been tested for special early or late failure rates, and it has been found that the early rates are not much different from the overall rate, computing the overall rate is easy. Just sum the number of failures to get a total, sum the number of survivals into each year to get a total, and divide. In the example, there were 217 system years of operation and 16 failures. That is an overall failure rate of 16/217 or 7.4%. In other words, expect 74 failures for every 1,000 systems each year. This is the number needed for the simulation model, to compute future failures. Compute the rate separately for each cohort.

STEP 4: SIMULATING FUTURE FAILURES

Now, simply take the data and use a step-by-step process to estimate future failures, and the costs of those failures. For simplicity, this manual includes a computer program which does the simulation, but the process can easily be done on a calculator or with spreadsheet software.

Take a single cohort at the start of the planning period. In the first year, some systems will fail. Some new systems may also be constructed, and some may be abandoned. The systems which fail may be reconstructed onsite, but the reconstruction will use current codes for that type of site. At the end of the year, there will be a different number of systems in each cohort. Formally,

$$S_{i,t} = S_{i,t-1}(1-p_i) = N_{i,t} - A_{i,t} \qquad [8]$$

where

$S_{i,t}$ is the stock of systems in cohort i at the end of year t

p_i is the failure rate for cohort i

$N_{i,t}$ is the number of new systems constructed in year t in cohort

$A_{i,t}$ is the number of systems abandoned from cohort i in year t

If there are 1000 systems at the start in the cohort, a 1% failure rate, 50 systems constructed per year, and no abandonments, then the stock for this system will be as follows:

Start: 1000

year 1: 1000 - 10.0 + 50 - 0 = 1040

year 2: 1040 - 10.4 + 50 - 0 = 1080

year 3: 1080 - 10.8 + 10 - 0 = 1119

year 4: 1119 - 11.2 + 50 - 0 = 1158

year 5: [etc.]

This analysis is done for each cohort for twenty years, and then the results for each planning year are added up. Thus, if cohort A had 10 failures in year 1, cohort B might have 7 and cohort C might have 5. The total would be 22 failures in year 1. If each failure cost $2000 to fix, on average, the cost that year would be $44,000. Given an interest rate, standard techniques can be used to compute the present worth (discounted) cost of these future repairs.

There is one area where care is required, which is the assumption about reconstruction. The easiest option is to assume no reconstruction

-- that is, not to worry about future failures from repaired systems. If failure rates are small, this could be acceptable.

However, to consider those failures, each reconstructed system must be assigned to a cohort. For example, consider a failure for a system on sand, built between 1957-1972. One option would be to assume that the reconstructed system was just like the one which failed, and put it back in that cohort. Alternatively, assume that the system is on the same soil class, but is built up to current codes. In that case, failures in any one cohort may be rebuilt in any other. It is straightforward to put those reconstructions back in, but time consuming. The basic equation becomes:

$$S_{i,t} = S_{i,t-1} (1-F_i) + N_{i,t} + A_{i,t} + F_{i,j,t} \qquad [9]$$

where

$F_{i,j,t}$ is the number of failures in cohort j for year t which are rebuilt as cohort i. Add this for each value of j.

The computation is eased by assuming that these reconstructed failures can not fail again in the year of reconstruction.

Example Application

The BASIC computer program and an example are contained in Appendix A. Most of the program is input and output, and these vary in different computer systems, so some revision may be required. However, that adaptation should be fairly easy. The model could also be implemented easily with a spreadsheet program such as Visicalc or Multiplan. The case in the example is a typical, but somewhat complicated one, with six cohorts and a 20-year planning period:

Cohort 1: systems built up to 1960, on sand. There are 100 such systems now, and the failure rate is 2.5%.

Cohort 2: systems built 1961-1975, on sand (70 now, 1.3% failure rate)

Cohort 3: 1976 or later, sand, 70 systems, 0.9% failure rate

Cohort 4: up to 1960, clay, 80 systems, 3.9% failure rate

Cohort 5: 1961-1975, clay, 100 systems, 2.2% failure rate

Cohort 6: 1976 or later, clay, 40 systems, 1.1% failure rate

There is no sewering, and no abandonment. From 1981-1990, 10 systems per year will be built in each of cohorts 3 and 6. From 1991-2000, 5 systems will be built annually in cohort 3, and 20 in cohort 6 since the good sandy lots are being used up. All failures are rebuilt on the same soil type to current standards, so failures are rebuilt as new members of cohorts 3 or 6.

Table 9 shows the results of this analysis, assuming $2500 per reconstruction in 1980 dollars, and a 7% real discount rate. The number of systems rises, and the average failure rate drops, since there are more systems in the cohorts with lower failure rates. The present value for repair of all these systems is $270,207, or $21.07 per system per year.

Table 10 shows the initial and final number of systems in each cohort. The numbers in cohorts 1, 2, 4, and 5 have dropped, since some fail and none are built. All the growth comes in cohorts 3 and 6, which gain the new construction and the reconstruction of all failures. Since these cohorts have lower failure rates, the average failure rate drops.

TABLE 9. EXAMPLE PROBLEM TABULATION

YEAR	STARTING COUNT	NEW SYSTEMS	FAILURES	ENDING COUNT	FAILURE	REPAIR COST	CUMULATIVE PRESENT WORTH
1981	460	20	9.80	480	2.13	$24500	$22897
1982	480	20	9.84	500	2.05	$24612	$44394
1983	500	20	9.89	520	1.98	$24736	$64587
1984	520	20	9.95	540	1.91	$24873	$83562
1985	540	20	10.01	560	1.85	$25021	$101402
1986	560	20	10.07	580	1.80	$25180	$118181
1987	580	20	10.14	600	1.75	$25351	$133968
1988	600	20	10.21	620	1.70	$25531	$148827
1989	620	20	10.29	640	1.66	$25722	$162818
1990	640	20	10.37	660	1.62	$25923	$175996
1991	660	25	10.45	685	1.58	$26133	$188412
1992	685	25	10.61	710	1.55	$26515	$200185
1993	710	25	10.76	735	1.52	$26905	$211350
1994	735	25	10.92	760	1.49	$27304	$221939
1995	760	25	11.08	785	1.46	$27712	$231983
1996	785	25	11.25	810	1.43	$28127	$241510
1997	810	25	11.42	835	1.41	$28550	$250549
1998	835	25	11.59	860	1.39	$28981	$259123
1999	860	25	11.77	885	1.37	$29418	$267257
2000	885	25	11.95	910	1.35	$29863	$274974

TABLE 10. OVERALL RESULTS OF EXAMPLE PROBLEM

Cohort #	Years	Soil	Number in 1980	Number in 2000	% change
1	to 1960	sand	100	60	-40%
2	1961-1975	sand	70	54	-23%
3	1976 or later	sand	70	276	+266%
4	to 1960	clay	80	36	-55%
5	1961-1975	clay	100	64	-36%
6	1976 or later	clay	40	420	+950%

REFERENCES

1. Hill, D.E. and C.R. Frink, "Septic System Longevity Increased by Improved Design." Journal WPCF 52(8):2199-2203, August 1980. and

   Hill, D.E. and C.R. Frink, "Longevity of Septic Systems in Connecticut Soils," Bulletin 747, Connecticut Agricultural Experiment Station, New Haven, June 1974. [and WPCF, 80]

2. Clayton, J.W. "A Study of Septic Tank Survival Data 1952-1974 in Fairfax County, Virginia", Journal of Environmental Health, 36:562-567, 1974. and

   Bowman, J.O. "A Study of Septic Tank Soil Absorption Systems," in N.I. McClelland and J.L. Evans (ed.), Individual Onsite Wastewater Systems: Proceedings of the Seventh National Conference, 1980. Ann Arbor: National Sanitation Foundation, 1981. pp. 109-114. and

   Bowman, J.O. "Reliability of Onsite Sewage Disposal Systems in Fairfax County, Virginia", Journal of Environmental Health, 44(5): 249-252, March-April, 1982.

3. Saxton, G.B. and J.M. Zeneski. "Prediction of Septic-System Failures," Journal of the Environmental Engineering Division, Proceedings ASCE 105(EE3):503-509, June 1979.

4. Department of County Engineer-Facilities, Los Angeles, California. Waste Treatment Management for the Malibu/Topangea Area (Task 2-1D, Phase II: SCAG 208 Areawide Waste Treatment Management Planning, June 1978.

5. Appendix: Study of Seepage Pits. Public Health Service, Cincinnati, 1961.

6. Bendixen, T.W., A.A. McMahan, R.E. Thomas, and J.B. Coulter. Effects of Food Waste Grinders on Septic Tank Systems, Public Health Service, Cincinnati, 1961.

7. Coulter, J.B., T.W. Bendixen, and A.B. Edwards. Practical Design Criteria for Seepage Beds. Public Health Services, Cincinnati, 1960.

8. Coulter, J.B. and T.W. Bendixen. Effectiveness of the Distribution Box, Public Health Services, Cincinnati, 1958.

9. DeWalle, F.B. "Failure Analysis of Septic Tank Systems," Unpublished, n.d.

10. Sullivan, G.M., J.B. Coulter, and T.W. Bendixen. A Study of Senial

Distribution for Soil Absorption Systems. Public Health Service, Cincinnati, 1959.

11. Weak, Victor D. A Technology Assessment Methodology, Volume VI Water Pollution: Domestic Wastes. MITRE Corp., June 1971. PB 202-778-6.

12. Winneberger, J.T. "Sanitary Surveys and Survival Curves of Septic Tank Systems." Journal of Environmental Health, 38(1):36-39, July 1975.

APPENDIX A

SAMPLE ANALYSIS

```
Run Date is 821018.  Time is 164430.  Operator is Hudson

Cost per repair is  $2500.  Real interest rate is   7.0%

Starting Year is 1980.  Number of cohorts is  6.  Number of years is 20.


Cohort  1 covers 1900 to 1960 with physical condition:Sand
For cohort 1, there are    100 starting systems, and p[failure]= 2.50

Cohort  2 covers 1961 to 1975 with physical condition:Sand
For cohort 2, there are     70 starting systems, and p[failure]= 1.30

Cohort  3 covers 1976 to 2000 with physical condition:Sand
For cohort 3, there are     70 starting systems, and p[failure]= 0.90

Cohort  4 covers 1900 to 1960 with physical condition:Clay
For cohort 4, there are     80 starting systems, and p[failure]= 3.90

Cohort  5 covers 1961 to 1975 with physical condition:Clay
For cohort 5, there are    100 starting systems, and p[failure]= 2.20

Cohort  6 covers 1976 to 2000 with physical condition:Clay
For cohort 6, there are     40 starting systems, and p[failure]= 1.10



Cohort  3; year 1981, net new systems =      10
Cohort  3; year 1982, net new systems =      10
Cohort  3; year 1983, net new systems =      10
Cohort  3; year 1984, net new systems =      10
Cohort  3; year 1985, net new systems =      10
Cohort  3; year 1986, net new systems =      10
Cohort  3; year 1987, net new systems =      10
Cohort  3; year 1988, net new systems =      10
Cohort  3; year 1989, net new systems =      10
Cohort  3; year 1990, net new systems =      10
Cohort  3; year 1991, net new systems =       5
Cohort  3; year 1992, net new systems =       5
Cohort  3; year 1993, net new systems =       5
Cohort  3; year 1994, net new systems =       5
Cohort  3; year 1995, net new systems =       5
Cohort  3; year 1996, net new systems =       5
Cohort  3; year 1997, net new systems =       5
Cohort  3; year 1998, net new systems =       5
Cohort  3; year 1999, net new systems =       5
Cohort  3; year 2000, net new systems =       5
```


BEGIN FIRST LINE OF TEXT HERE
CENTER OF PAGE
TOP OF IMAGE AREA
UNDERLINED HEADINGS SHOULD BE IN CAPITAL LETTERS
BEGIN LAST LINE OF TEXT
3/8"
BOTTOM OF IMAGE AREA. OUTSIDE DIMENSION FOR TABLES AND ILLUSTRATIONS
PAGE NUMBER
EPA Form 2350-4 (4-80)
(PREVIOUSLY EPA FORM 330)
TYPING GUIDE SHEET

```
Cohort  6; year 1981, net new systems =      10
Cohort  6; year 1982, net new systems =      10
Cohort  6; year 1983, net new systems =      10
Cohort  6; year 1984, net new systems =      10
Cohort  6; year 1985, net new systems =      10
Cohort  6; year 1986, net new systems =      10
Cohort  6; year 1987, net new systems =      10
Cohort  6; year 1988, net new systems =      10
Cohort  6; year 1989, net new systems =      10
Cohort  6; year 1990, net new systems =      10
Cohort  6; year 1991, net new systems =      20
Cohort  6; year 1992, net new systems =      20
Cohort  6; year 1993, net new systems =      20
Cohort  6; year 1994, net new systems =      20
Cohort  6; year 1995, net new systems =      20
Cohort  6; year 1996, net new systems =      20
Cohort  6; year 1997, net new systems =      20
Cohort  6; year 1998, net new systems =      20
Cohort  6; year 1999, net new systems =      20
Cohort  6; year 2000, net new systems =      20

Rebuild failures in cohort  1 as new members of cohort  3
Rebuild failures in cohort  2 as new members of cohort  3
Rebuild failures in cohort  3 as new members of cohort  3
Rebuild failures in cohort  4 as new members of cohort  6
Rebuild failures in cohort  5 as new members of cohort  6
Rebuild failures in cohort  6 as new members of cohort  6
```

YEAR	STARTING COUNT	NEW SYSTEMS	FAILURES	ENDING COUNT	% FAILURE	REPAIR COST	CUM. PRES. VALUE
1981	460	20	9.80	480	2.13	$24500	$22897
1982	480	20	9.84	500	2.05	$24612	$44394
1983	500	20	9.89	520	1.98	$24736	$64587
1984	520	20	9.95	540	1.91	$24873	$83562
1985	540	20	10.01	560	1.85	$25021	$101402
1986	560	20	10.07	580	1.80	$25180	$118181
1987	580	20	10.14	600	1.75	$25351	$133968
1988	600	20	10.21	620	1.70	$25531	$148827
1989	620	20	10.29	640	1.66	$25722	$162818
1990	640	20	10.37	660	1.62	$25923	$175996
1991	660	25	10.45	685	1.58	$26133	$188412
1992	685	25	10.61	710	1.55	$26515	$200185
1993	710	25	10.76	735	1.52	$26905	$211350
1994	735	25	10.92	760	1.49	$27304	$221939
1995	760	25	11.08	785	1.46	$27712	$231983
1996	785	25	11.25	810	1.43	$28127	$241510
1997	810	25	11.42	835	1.41	$28550	$250549
1998	835	25	11.59	860	1.39	$28981	$259123
1999	860	25	11.77	885	1.37	$29418	$267257
2000	885	25	11.95	910	1.35	$29863	$274974

```
Cohort  1 covers 1900 to 1960 with physical condition:Sand
For cohort  1, final count is      60

Cohort  2 covers 1961 to 1975 with physical condition:Sand
For cohort  2, final count is      54

Cohort  3 covers 1976 to 2000 with physical condition:Sand
For cohort  3, final count is     276

Cohort  4 covers 1900 to 1960 with physical condition:Clay
For cohort  4, final count is      36

Cohort  5 covers 1961 to 1975 with physical condition:Clay
For cohort  5, final count is      64

Cohort  6 covers 1976 to 2000 with physical condition:Clay
For cohort  6, final count is     420
```


EPA Form 2350-4 (4-80)
(PREVIOUSLY CIN: EPA FORM 267)
TYPING GUIDE SHEET

```
0010 Rem     THIS IS A PROGRAM FOR SIMULATING COHORT SURVIVAL OF SEPTIC SYSTE
MS
0020 Rem     Version 1, written by J. Hudson, USR&E, 10/17/81
0030 input "unit # is ",z$
0035 z$="U0"&z$
0040 open print #1,QUEUE=z$
0045 select print #1,lmargin=6,rmargin=96,lines=50
0100 dim s0(10),p(10),n(10,20),s(10,20),f(10,20),f1(20),s1(20),n1(20),r(10),
a$(10),a1$(10),a2$(10)
0105 Rem
0110 Rem     GENERAL INPUT
0120 Rem
0122 input "Operator is (up to 8 characters) ",q$
0124 input "Repair cost in constant dollars is $",c
0126 input "Real interest rate in % per year is ",d
0130 Input "Number of cohorts (max is 10) is ",k
0140 Input "Starting year is 19",y
0150 y=1900+y
0160 input "Number of years in planning period (max is 20) is ",m
0166 printusing #1, 9900,date,time,q$,skip(2)
0168 printusing #1, 9910,c,d,skip(2)
0170 printusing #1,9800,y,k,m,skip(3)
0175 d=d/100
0190 print page
0200 printusing 9000,y,skip(2)
0205 for i=1 to k
0210    printusing 9100,i
0212    input "first year of construction included ",a$(i)
0214    input "last year of construction included ",a1$(i)
0215    input "physical characteristics included (max 16 char)",a2$(i)
0220    input "starting number of systems is ",S0(i)
0230    input "failure probability in percent per year is ",p(i)
0232    p(i)=p(i)/100
0240    if p(i)>=1.0 or if p(i)<0 then go to 9999
0243 printusing #1, 9830,i,a$(i),a1$(i),a2$(i)
0245 printusing #1,9810,i,s0(i),p(i)*100.0+.005,skip(2)
0250 next i
0300 printusing 9200,skip(2)
0305 print #1,skip(3)
0310 for i=1 to k
0315 if convert(a1$(i)) y then goto 0375:error goto 0320
0320    printusing 9100,i
0330    for j=1 to m
0340       y1=y+j
0350       printusing 9300,y1
0360       input "Net new systems total ",N(i,j)
0365      if N(i,j)<> 0 then printusing #1,9820,i,y1,N(i,j)
0370    next j
0375 print #1,skip(1)
0380 next i
0390 input "Rebuild failures? (Y or N) ",R$
0400 if r$<>"Y" then go to 0490
0410 for i=1 to k
0420    printusing 9700,i
0430    input "rebuild as a member of cohort ",r(i)
```


EPA Form 2350-4 (4-80)
(PREVIOUSLY CIN, EPA FORM 2571)
TYPING GUIDE SHEET

```
0440    printusing #1,9710,i,r(i)
0450  next i
0490 rem
0492 rem      PRINT TABLE HEADING
0494 rem
0496 print #1,page
0498 printusing #1,9410
0500 printusing #1,9400,skip(2)
0502 rem
0504 rem        COMPUTE INITIAL YEAR SPECIALLY
0506 rem
0510 for i=1 to k
0520     f(i,1)=s0(i)*p(i)
0530     f1(1)=f1(1)+f(i,1)
0540     s(i,1)=s0(i)-f(i,1)+n(i,1)
0550     n1(1)=n1(1)+n(i,1)
0560     s1(1)=s1(1)+s(i,1)
0565     s9=s9+s0(i)
0568  next i
0570 for i=1 to k
0572    s(r(i),1)=s(r(i),1)+f(i,1)
0574    s1(1)=s1(1)+f(i,1)
0576 next i
0582  c1=f1(1)*c
0592  c9=c1/(1+d)
0600     printusing #1,9500,y+1,s9+.5,n1(1),f1(1)+.005,s1(1)+.5,100*f1(1)/s9
+.005,c1+.5,c9+.5
0990 rem
0992 rem        COMPUTE AND PRINT YEARS 2-20
0994 rem
1000 for j=2 to m
1010     y1=y+j
1020     for i=1 to k
1030        f(i,j)=s(i,j-1)*p(i)
1040        f1(j)=f1(j)+f(i,j)
1050        s(i,j)=s(i,j-1)-f(i,j)+n(i,j)
1060        n1(j)=n1(j)+n(i,j)
1070        s1(j)=s1(j)+s(i,j)
1080     next i
1090 if r$<>"Y" then go to 1500
1100 for i=1 to k
1110    s(r(i),j)=s(r(i),j)+f(i,j)
1120    s1(j)=s1(j)+f(i,j)
1130 next i
1500     c1=f1(j)*c
1510     c9=c9+c1/((1+d)j)
1520     printusing #1,9500,y1,s1(j-1)+.5,n1(j),f1(j)+.005,s1(j)+.5,100*f1(j
)/s1(j-1)+.005,c1+.5,c9+.5
1530  next j
1540 rem
1550 rem      PRINTOUT COHORT STOCKS AT END
1560 rem
1570 print #1,page
1580  for i=1 to k
1585    printusing #1,9830,i,a$(i),a1$(i),a2$(i)
```

```
1590      printusing #1,9600,i,s(i,m)+.5,skip(2)
1600   next i
9000 %input failure probability and (####) number of systems for each cohort
##
9100 %cohort (##)
9200 %For each cohort and each year, input the net number of new systems@@
9300 %year is (####)
9400%YEAR     COUNT     SYSTEMS    FAILURES     COUNT    FAILURE    COST       VALU
E@@
9410%       STARTING      NEW                  ENDING       %%      REPAIR    CUM.
PRES.
9500%(####)    (######)  (-######)    (######.##)  (######)     (##.##)    ($###
####)    ($########)
9600 %For cohort (##), final count is (######)@@
9700%Failures in cohort (##)
9710 %Rebuild failures in cohort (##) as new members of cohort (##)
9800%Starting Year is (####).  Number of cohorts is (##).  Number of years i
s (##).@@
9810%For cohort (#), there are (######) starting systems, and p[failure]=(##
.##)@@
9820 %Cohort (##); year (####), net new systems = (-######)
9830 %Cohort (##) covers (####) to (####) with physical condition:(##########
########)
9900 %Run Date is (######).  Time is (######).  Operator is (########)@@
9910 %Cost per repair is ($#####).  Real interest rate is (##.#)%%@@
9999 end
```

APPENDIX B

TABLE OF STUDENT'S t - DISTRIBUTION

TABLE OF STUDENT'S t - DISTRIBUTION

Degrees of Freedom	0.05
1	12.706
2	4,303
3	3.182
4	2.776
5	2.571
6	2.447
7	2.365
8	2.306
9	2.262
10	2.228
11	2.201
12	2.179
13	2.160
14	2.145
15	2.131
16	2.120
17	2.110
18	2.101
19	2.093
20	2.086
21	2.080
22	2.074
23	2.069
24	2.064
25	2.060
26	2.056
27	2.052
28	2.048
29	2.045
30	2.042
∞	1.95996

Reproduced from Statistical Methods for Research Workers, 6th ed., with the permission of the author, R.A. Fisher, and his publisher, Oliver and Boyd, Edinburgh.